Editorial

Introduction to the special issue of JWS with selected papers from ISWC 2005

The series of International Semantic Web Conferences (ISWC) started in 2002 and over the years has firmly established itself as the premiere international event in this rapidly expanding field. This special issue contains four of the best papers submitted to the Research/Academic Track of ISWC 2005, held in Galway, Ireland, on November 6–10, 2005. ISWC 2005 attracted several hundred participants from academia, government and industry, thus reflecting the growing interest in all sectors of society towards semantic web technologies.

The Research/Academic Track of ISWC 2005 had 217 submissions, an increase over the number of papers submitted to ISWC 2004. This confirmed the robustness of the research base in this area, at a time of increased competition from other events. In total 54 papers were accepted, for a 25% overall acceptance rate.

The four papers selected for this special issue illustrate the range of research areas tackled by the semantic web community. The work by Peter Mika brings together a number of interesting strands of research—folksonomies, social network analysis, ontologies and text mining, in an attempt to go beyond the traditional dichotomy between formally axiomatized ontologies and informal social tagging, and try and achieve an initial synthesis. An interesting (and possibly controversial) element of Peter Mika’s analysis is his view that folksonomies and ontologies are actually different points over a common spectrum of increased formalization and systematization of categorical systems. While a number of people may disagree with this view, an important element of Peter Mika’s work is the instantiation of these ideas in concrete applications, which bring together tagging systems, web mining technology and ontologies. His paper in this special issue includes two case studies. In the first one he shows the emergence of a lightweight ontology from an analysis of tags in the popular del.icio.us system. In the second one he presents his Flink system, which provides a visualization of the social networks and research interests of researchers in the semantic web. This information is extracted from web pages using text mining methods and presented uniformly in terms of a common ontology. In addition, techniques drawn from social science research are used to derive formal measures about the social network.

The paper by David Huynh, Stefano Mazzocchi, and David Karger tackles a crucial issue for our community. Although the size and use of the semantic web continues to increase, it is still dwarfed by the size and use of the original web. This paper presents PiggyBank, a creative approach to grow the semantic web by allowing users of the regular web to contribute semantic markup with minimal effort. By enabling users to annotate web pages as they access them, PiggyBank collects semantic annotations that describe and classify their content according to that particular user’s use of the information. Users mark pages by tagging them much as they would do in del.icio.us, and PiggyBank extracts additional markup through screenscrapers and by retrieving RDF or RSS markup provided by the original publishers of the page. An interesting aspect of PiggyBank is that in addition to growing folksonomies (folk taxonomies) by collecting tags, it can also acquire relations among the tags from the user to create what the authors call folksologies (folk ontologies). Each user’s annotations are then published in shared Semantic Banks, and include provenance information regarding which users, screenscrapers, or publishers contributed the markup. If this approach were to become routine practice of Web users, these shared Semantic Banks would quickly populate the semantic web.

The work presented by Simon Miles, Sylvia Wong, Weijian Fang, Paul Groth, Klaus-Peter Zauner, and Luc Moreau is representative of the integration of semantic web technologies in e-science. Complex on-line analysis of scientific data often involves complex integration and processing of distributed data. When these “in silico” experiments are performed, the results must be validated by ensuring that the data analysis processes were correctly used and executed. This paper nicely integrates scalable techniques to capture this kind of data provenance trails and semantic web techniques for describing with rich metadata the details of the data sources and computations. The paper describes several use cases of the use of provenance information by scientists. The scientific community may be one of the earliest adaptors and beneficiaries of semantic web technologies.

The paper by Peter Plessers, Olga De Troyer and Sven Casteleyn addresses another key issue for the semantic web, that of developing effective methods to support Ontology Evolution.
The reason why this topic is of particular importance in the context of the semantic web is that, given its open and distributed nature, it is simply not possible to control the evolution of semantic information, in the same way that it was possible in the days of ‘closed’ knowledge based systems, where both knowledge acquisition and system evolution where tightly controlled by a team of developers. Hence, in the semantic web context it is very important to have effective methods for managing ontology evolution, as changes in one ontology may impact on a large number of applications based on that particular ontology. Plessers et al. provide a methodological framework, a specification language, and change detection algorithms to support ontology evolution. Thus, an important element of their approach is that they provide both a mechanism for developers to describe the changes they have made to an ontology, according to different viewpoints, and also ways to propagate changes and maintain consistency. The authors also describe a prototype implementation of their system, developed as a plug-in for the Protégé ontology editor.

In summary, this special issue illustrates some crucial and intriguing new directions in semantic web research. By exploring approaches that emphasize the combination of folksonomies and ontologies, rather than presenting them as diametrical choices, the research community is addressing the crucial issue of scaling up the population and use of the semantic web. By addressing the ontology lifecycle and methodology, researchers are proactively anticipating and resolving the issues crucial to supporting long-term large distributed communities using the semantic web. Through approaches that integrate semantic web technologies into e-science environments, researchers continue to demonstrate the benefits of these technologies to a vibrant scientific community that presents many challenges and rewarding possibilities. We invite you to learn more about these exciting topics and read this special issue as well as the proceedings of ISWC 2005. Our young field continues to surprise us with creative approaches to the great challenges we face.

Yolanda Gil *

Information Sciences Institute, University of Southern California, 4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90292, USA

Enrico Motta

Knowledge Media Institute, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK

* Corresponding author.

E-mail addresses: gil@isi.edu (Y. Gil), e.motta@open.ac.uk (E. Motta)

12 December 2006

Available online 22 January 2007